SIAM J. APPL. MATH.
Vol. 68, No. 6, pp. 1618–1640
© 2008 Society for Industrial and Applied Mathematics

EXISTENCE, UNIQUENESS, AND A CONSTRUCTIVE SOLUTION ALGORITHM FOR A CLASS OF FINITE MARKOV MOMENT PROBLEMS*

LAURENT GOSSE† AND OLOF RUNBORG‡

Abstract. We consider a class of finite Markov moment problems with an arbitrary number of positive and negative branches. We show criteria for the existence and uniqueness of solutions, and we characterize in detail the nonunique solution families. Moreover, we present a constructive algorithm to solve the moment problems numerically and prove that the algorithm computes the right solution.

Key words. finite Markov moment problem, inverse problems, exponential transform

AMS subject classifications. 30E05, 15A29, 65H10, 65M99

DOI. 10.1137/070692510

1. Introduction.

We aim at inverting a moment system often associated with the prestigious name of Markov. The original form of the problem is the following. Given a *finite set of moments* m_k for $k = 1, \dots, K$, find a bounded measurable *density* function f satisfying

$$m_k = \int_{\mathbb{R}} x^{k-1} f(x)dx, \qquad 0 \le f \le 1, \qquad k = 1, \dots, K. \tag{1.1}$$

The condition for the existence of solutions $f(x)$ to this problem is classical [1, 2]. In general, solutions are not unique, unless more conditions are given, e.g., based on entropy minimization [3, 4] or L^∞-minimization [19, 18]. A typical result is that the unique solution for even K is piecewise constant, taking values in $\{0, 1\}$. More precisely, if $K = 2n$, then f is of the form

$$f(x) = \sum_{j=1}^{n} \chi_{[y_i, x_i]}(x), \tag{1.2}$$

where $\chi_I(x)$ is the characteristic function for the interval I and

$$y_1 < x_1 < y_2 < x_2 < \cdots < y_n < x_n. \tag{1.3}$$

See Theorem 6.1 and consult, e.g., [5, 8, 17, 23, 25] for general background on moment problems.

A reduced form of the finite moment problem is to search for solutions to (1.1) which are precisely of the form (1.2), (1.3). One then obtains an algebraic problem for the branch values,

$$m_k = \frac{1}{k} \sum_{j=1}^{n} x_j^k - y_j^k, \qquad k = 1, \dots, K = 2n. \tag{1.4}$$

*Received by the editors May 20, 2007; accepted for publication (in revised form) March 10, 2008; published electronically June 6, 2008.
http://www.siam.org/journals/siap/68-6/69251.html

†IAC–CNR "Mauro Picone" (sezione di Bari), Via Amendola 122/D, 70126 Bari, Italy (l.gosse@ba.iac.cnr.it).
‡Department of Numerical Analysis, CSC, KTH, 10044 Stockholm, Sweden (olofr@nada.kth.se).

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Finding $\{x_j\}$ and $\{y_j\}$ from $\{m_k\}$ is an ill-conditioned problem when the branch values of the solution come close to each other; the Jacobian of the problem is a Vandermonde matrix, and iterative numerical resolution routines require extremely good starting guesses when the matrix degenerates. For less than four moments, a direct method based on solving polynomial equations was presented in [21]. Routines based on the simplex algorithm were proposed in [19]. Another algorithm was presented by Koborov, Sklyar, and Fardigola in [16, 24] in the slightly modified setting where f takes values in $\{-1, 1\}$ instead of $\{0, 1\}$. It consists of solving a sequence of high degree polynomial equations, constructed through a rather intricate process with unclear stability properties. In [14] we showed that this algorithm can be drastically simplified and adapted to (1.4). Later, in [15], we also gave a direct proof that the simplified algorithm indeed computes the correct solution, relying on the classical Newton identities and Toeplitz matrix theory.

The moment problem has many applications in, for instance, probability and statistics [10, 7] but also in areas like wave modulation [6, 22] and "shape from moments" inverse problems [11]. Our own motivation comes from a quite different field, namely, multiphase geometrical optics [3, 4, 12, 13, 14, 21]. In this application one needs to solve a system of nonlinear hyperbolic conservation laws. To evaluate the flux function in the PDEs a system like (1.4) must be solved. In a finite difference method this means that the system must be inverted once for every point in the computational grid repeatedly in every timestep. It is thus important that the inversion can be done quickly and accurately; this difficulty has been a bottleneck in computations. In [14] we used the simplified algorithm mentioned above for numerical implementation inside a shock-capturing finite difference solver. It is our aim here to develop better algorithms and understanding to open the way for the processing of intricate wave-fields with large K and thus complement the seminal paper [4], where the multiphase geometrical optics PDEs were first proposed.

In this paper we are concerned with a generalization of (1.4). In the geometrical optics application, the number of moments K is typically not even, and one can have a variable number of positive (x_k) and negative (y_k) branches. We thus consider the problem

$$m_k = \sum_{j=1}^{n_x} x_j^k - \sum_{j=1}^{n_y} y_j^k, \qquad k = 1, \ldots, K, \tag{1.5}$$

where $n_x + n_y = K$ but where n_x and n_y are not necessarily equal. We study existence and uniqueness of solutions to this problem (Theorem 4.1). In particular, we are interested in how and when uniqueness is lost. For these cases we characterize the family of solutions that exists. The reason is to understand what happens numerically close to degenerate solutions, which is an important feature in the application we have in mind: In the exact solution to the multiphase geometrical optics PDEs, the moment problem is typically degenerate for large domains; the numerical approximation is almost degenerate.

We also give constructive algorithms to solve (1.5) and prove that they generate the right solution (Theorem 2.1). In a future paper we will study the numerical stability of these algorithms. Experimentally we note, for instance, that to compute the next moment, Algorithm 3 is much more stable than Algorithm 1. The difficulty lies in understanding perturbations around degenerate solutions, which is where the algorithms are most unstable. For this the insights of this paper will be of importance.

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

REMARK 1. *The problem* (1.5) *can be cast in the form of* (1.1) *if one demands that the density function* $f(x)$ *be of the form*

$$f(x) = \sum_{j=1}^{n_x} \operatorname{sgn}(x_j)\,[H(x) - H(x - |x_j|)] - \sum_{j=1}^{n_y} \operatorname{sgn}(y_j)\,[H(x) - H(x - |y_j|)]\,, \tag{1.6}$$

and we rescale the moments $m_k \to k m_k$. *For the case* $n_x = n_y = n$ *and* $K = 2n$ *with interlaced branch values* (1.3), *this reduces to* (1.2).

This paper is organized as follows. In section 2 we present the algorithms for solving (1.5). Notation and various ways of describing a solution are subsequently introduced in section 3. Next we derive conditions for existence and uniqueness of solutions in section 4 and also discuss various properties of the solution, particularly when it is not unique. A theorem proving the correctness of the algorithms is proved in section 5. Finally, in section 6, we give additional properties of the elements of our algorithms and use these to relate our results back to the classical Markov theory.

2. Algorithms.

In this section we detail the algorithms that we propose for solving (1.5). The solution that we obtain is what we call the *minimal degree solution*, meaning that when the solution is not unique as many branch values as possible are zero. See section 4 for a precise definition. The algorithms are as follows; they may fail in case there is no solution to (1.5).

ALGORITHM 1 (computing $\{x_j\}$ and $\{y_j\}$).

1. *Construct the sequence* $\{a_k\}$ *as follows. Set* $a_0 = 1$ *and* $a_k = 0$ *for* $k < 0$. *For* $1 \le k \le K$, *let the elements be given as the solution to*

$$\begin{pmatrix} 1 & & & \\ -m_1 & 2 & & \\ \vdots & \ddots & \ddots & \\ -m_{K-1} & \dots & -m_1 & K \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_K \end{pmatrix} = \begin{pmatrix} m_1 \\ m_2 \\ \vdots \\ m_K \end{pmatrix}. \tag{2.1}$$

2. *Construct the matrix* $A_1 \in \mathbb{R}^{n_x \times n_x}$ *as*

$$A_1 = \begin{pmatrix} a_{n_y} & a_{n_y-1} & \cdots & a_{n_y-n_x+1} \\ a_{n_y+1} & a_{n_y} & \cdots & a_{n_y-n_x+2} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n_y+n_x-1} & a_{n_y+n_x-2} & \cdots & a_{n_y} \end{pmatrix}.$$

   Compute the rank of A_1. *Let* $\tilde{n}_x = \operatorname{rank} A_1$ *and* $\tilde{n}_y = n_y - n_x + \tilde{n}_x$.

3. *Construct the matrices* $\tilde{A}_0, \tilde{A}_1 \in \mathbb{R}^{\tilde{n}_x \times \tilde{n}_x}$ *as*

$$\tilde{A}_0 = \begin{pmatrix} a_{\tilde{n}_y+1} & a_{\tilde{n}_y} & \cdots & a_{\tilde{n}_y-\tilde{n}_x+2} \\ a_{\tilde{n}_y+2} & a_{\tilde{n}_y+1} & \cdots & a_{\tilde{n}_y-\tilde{n}_x+3} \\ \vdots & \vdots & \ddots & \vdots \\ a_{\tilde{n}_y+\tilde{n}_x} & a_{\tilde{n}_y+\tilde{n}_x-1} & \cdots & a_{\tilde{n}_y+1} \end{pmatrix},$$

$$\tilde{A}_1 = \begin{pmatrix} a_{\tilde{n}_y} & a_{\tilde{n}_y-1} & \cdots & a_{\tilde{n}_y-\tilde{n}_x+1} \\ a_{\tilde{n}_y+1} & a_{\tilde{n}_y} & \cdots & a_{\tilde{n}_y-\tilde{n}_x+1+1} \\ \vdots & \vdots & \ddots & \vdots \\ a_{\tilde{n}_y+\tilde{n}_x-1} & a_{\tilde{n}_y+\tilde{n}_x-2} & \cdots & a_{\tilde{n}_y} \end{pmatrix}.$$

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

4. *Solve the generalized eigenvalue problem*

$$\tilde{A}_0 \boldsymbol{v} = x \tilde{A}_1 \boldsymbol{v} \tag{2.2}$$

to get the $\{x_j\}$ *values of the minimal degree solution to* (1.5).

5. *To compute the* $\{y_j\}$ *values, the same process is used with* m_k *replaced by* $-m_k$ *and the roles of* n_x *and* n_y *interchanged.*

An alternative to Algorithm 1 is as follows.

ALGORITHM 2 (computing $\{x_j\}$ and $\{y_j\}$).

1. *Construct the matrices* $\tilde{A}_0$ *and* $\tilde{A}_1$ *as in steps* 1–3 *in Algorithm* 1.
2. *Denote the first column vector in* $\tilde{A}_0$ *by* $\tilde{\boldsymbol{a}}_0$ *by and solve*

$$\tilde{A}_1 \boldsymbol{c}' = -\tilde{\boldsymbol{a}}_0, \qquad \boldsymbol{c}' = (c_1, c_2, \ldots, c_{\tilde{n}_x})^T. \tag{2.3}$$

3. *Construct the polynomial*

$$P(z) = c_{\tilde{n}_x} + c_{\tilde{n}_x - 1} z + \cdots + c_1 z^{\tilde{n}_x - 1} + z^{\tilde{n}_x}.$$

The roots of $P(z)$ *are the* $\{x_j\}$ *values of the minimal degree solution to* (1.5) *(possibly together with some zeros).*

4. *To compute the* $\{y_j\}$ *values, the same process is used with* m_k *replaced by* $-m_k$ *and the roles of* n_x *and* n_y *interchanged.*

REMARK 2. *We note that the values of* a_k *in the definition* (2.1) *are independent of* K*, since the system matrix is triangular. We therefore consider the sequence without reference to* K *in any other respect than the fact that we are only able to compute elements with* $k \leq K$ *when we are given* K *moments. The largest index of the* a_k*-sequence appearing in the matrix* A_1 *is* $n_y + n_x - 1 < K$*. In the matrices* $\tilde{A}_0, \tilde{A}_1$ *it is* $\tilde{n}_y + \tilde{n}_x = n_y - n_x + 2\tilde{n}_x \leq n_y + n_x = K$*. Hence all three matrices can be constructed from the first* K *moments. Some properties of the* A_1 *matrix are detailed in section* 6.

Sometimes one is not interested in finding the individual $\{x_j\}$ and $\{y_j\}$ branch values but just wants the higher moments, defined as

$$m_k = \sum_{j=1}^{n_x} x_j^k - \sum_{j=1}^{n_y} y_j^k, \tag{2.4}$$

but now for $k > K$, *given* a solution $\{x_j\} \cup \{y_j\}$ to (1.5). (That this is well defined is shown later in Theorem 4.1.) For this case there is another algorithm which has empirically proven to be more stable than first computing $\{x_j\}$ and $\{y_j\}$ from Algorithm 1 or 2 and then entering the values into (2.4). We stress that this is precisely what is needed in order to compute K-multivalued solutions of the inviscid Burgers equation in geometrical optics, following the ideas of [4].

ALGORITHM 3 (computing m_{K+1}).

1. *Construct the* A_1 *matrix as in steps* 1–2 *of Algorithm* 1.
2. *Let*

$$\boldsymbol{a}_0 = (a_{n_y+1}, a_{n_y+2}, \ldots, a_{n_y+n_x})^T \in \mathbb{R}^{n_x},$$

and let $\bar{\boldsymbol{c}} = (c_1, c_2, \ldots, c_{n_x})^T$ *be one solution to*

$$A_1 \bar{\boldsymbol{c}} = -\boldsymbol{a}_0. \tag{2.5}$$

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

3. *The next moment is given by*

$$m_{K+1} = -(K+1)\sum_{j=1}^{n_x} c_j a_{K+1-j} - \sum_{j=1}^{K} m_j a_{K+1-j}.$$

We recall that Algorithm 1 has been shown to be numerically efficient in the paper [14]. The justification of these algorithms is given in section 5, where we show the following theorem.

THEOREM 2.1. *If a solution to* (1.5) *exists, then the following hold.*

(i) *In Algorithm* 1, *the matrix* $\tilde{A}_1$ *is nonsingular. The generalized eigenvalue problem in* (2.2) *is well defined and the generalized eigenvalues (counting algebraic multiplicity) are the* $\{x_j\}$*-values of the minimal degree solution to* (1.5) *plus* $\tilde{n}_x - D_{\min}$ *zeros. (See* (3.7) *for the definition of* $D_{\min}$*.)*

(ii) *In Algorithm* 2, $\boldsymbol{c}'$ *is well defined,*

$$P(z) = \det(zI - \tilde{A}_1^{-1}\tilde{A}_0), \tag{2.6}$$

and the roots of $P(z)$ *are the* $\{x_j\}$*-values of the minimal degree solution to* (1.5) *plus* $\tilde{n}_x - D_{\min}$ *zeros.*

(iii) *In Algorithm* 3, *the computed moment satisfies*

$$m_{K+1} = \sum_{j=1}^{n_x} x_j^{K+1} - \sum_{j=1}^{n_y} y_j^{K+1}$$

for all solutions $\{x_j\} \cup \{y_j\}$ *to* (1.5).

We postpone the proof of Theorem 2.1 to section 5. We just note here that the last point in Algorithms 1 and 2 can easily be explained by the symmetry of the problem. Indeed, the negative of (1.5),

$$-m_k = \sum_{j=1}^{n_y} y_j^k - \sum_{j=1}^{n_x} x_j^k, \qquad k = 1, \dots, K,$$

is of the same form as (1.5) itself, with the roles of n_x, $\{x_j\}$ and n_y, $\{y_j\}$ interchanged.

3. Preliminaries. We will use three different ways of describing the solution to (1.5). First, we have a set of numbers $\{x_j\}_{j=1}^{n_x}$ and $\{y_j\}_{j=1}^{n_y}$, solving (1.5). We call those numbers branch values. Second, we have a pair of polynomials (p, q) of degrees at most n_x and n_y, respectively, in the z variable. Third, we have a pair of coefficient vectors $\boldsymbol{c} = (c_0, \dots, c_{n_x})^T \in \mathbb{R}^{n_x+1}$ and $\boldsymbol{d} = (d_0, \dots, d_{n_x})^T \in \mathbb{R}^{n_y+1}$. These three representations are related as

$$p(z) = (1 - x_1 z)\cdots(1 - x_{n_x} z) = c_0 + c_1 z + \cdots + c_{n_x-1} z^{n_x-1} + c_{n_x} z^{n_x} \tag{3.1}$$

and

$$q(z) = (1 - y_1 z)\cdots(1 - y_{n_y} z) = d_0 + d_1 z + \cdots + d_{n_y-1} z^{n_y-1} + d_{n_y} z^{n_y}. \tag{3.2}$$

It is clear that there is a one-to-one correspondence between these ways of describing the solution if we disregard the ambiguity in the ordering of the numbers $\{x_j\}$ and $\{y_j\}$. Generally, we will use the notation $\mathrm{Deg}(p)$ to denote the degree of a polynomial p, and, for a given coefficient vector $\boldsymbol{c}$, we systematically write P_c to denote the corresponding polynomial (3.1).

DEFINITION 3.1. *We call the pair of polynomials* (p, q) a (polynomial) solution *to* (1.5) *if the following hold.*

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

1. *The degrees of* p *and* q *are at most* n_x *and* n_y*:*

$$\operatorname{Deg}(p) \le n_x, \qquad \operatorname{Deg}(q) \le n_y. \tag{3.3}$$

2. *They are normalized to one at the origin:*

$$p(0) = q(0) = 1. \tag{3.4}$$

3. *Their roots* $\{\tilde{x}_j\}$ *and* $\{\tilde{y}_j\}$ *satisfy*

$$m_k = \sum_{j=1}^{\operatorname{Deg}(p)} \tilde{x}_j^{-k} - \sum_{j=1}^{\operatorname{Deg}(q)} \tilde{y}_j^{-k}, \qquad k = 1, \ldots, K. \tag{3.5}$$

We note that the roots cannot be zero because of (3.4).

Next, we have the following.

DEFINITION 3.2. *A pair of vectors*

$$\boldsymbol{c} = (c_0, \ldots, c_{n_x})^T \in \mathbb{R}^{n_x+1} \text{ and } \boldsymbol{d} = (d_0, \ldots, d_{n_y})^T \in \mathbb{R}^{n_y+1}$$

is said to be a (coefficient) solution *to* (1.5) *if the corresponding pair* (P_c, P_d) (3.1)–(3.2) *realizes a polynomial solution to* (1.5).

The number of branch values is always n_x and n_y, respectively. Some of them may be zero, and they do not need to be distinct. The number of nonzero branch values is $\operatorname{Deg}(p)$ and $\operatorname{Deg}(q)$, respectively. The degree of a solution can then also be defined.

DEFINITION 3.3. *The* degree *of a solution to* (1.5) *is the number of nonzero* x_j*-values. This number is equivalent to* $\operatorname{Deg}(p)$.

Given any polynomial pair satisfying (3.4), we say that it generates the moment sequence $\{m_k\}$ if m_k is given by (3.5) for all k. In turn, each sequence of moments $\{m_k\}$ generates the corresponding $\{a_k\}$ sequence through (2.1). We define the big matrix

$$A = \begin{pmatrix} a_{n_y+1} & a_{n_y} & \cdots & a_{n_y-n_x+1} \\ a_{n_y+2} & a_{n_y+1} & \cdots & a_{n_y-n_x+2} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n_y+n_x} & a_{n_y+n_x-1} & \cdots & a_{n_y} \end{pmatrix} \in \mathbb{R}^{n_x \times (n_x+1)}.$$

We let the columns of A be denoted $\boldsymbol{a}_0, \ldots, \boldsymbol{a}_{n_x}$, and we note that

$$A = \begin{pmatrix} | & & | \\ \boldsymbol{a}_0 & \cdots & \boldsymbol{a}_{n_x} \\ | & & | \end{pmatrix} = \begin{pmatrix} & & | \\ & A_0 & \boldsymbol{a}_{n_x} \\ & & | \end{pmatrix} = \begin{pmatrix} | & & \\ \boldsymbol{a}_0 & A_1 & \\ | & & \end{pmatrix}. \tag{3.6}$$

Hence A_0 and A_1 constitute the first and last n_x columns of A, respectively. When $\boldsymbol{a}_0 \in \operatorname{range} A_1$ and $\boldsymbol{a}_0 \neq 0$, let

$$D_{\min} = \operatorname{argmin}_{j>0} \boldsymbol{a}_0 \in \operatorname{span}\{\boldsymbol{a}_1, \ldots, \boldsymbol{a}_j\}, \tag{3.7}$$

and set $D_{\min} = 0$ if $\boldsymbol{a}_0 = 0$. Moreover, define

$$D_{\max} = D_{\min} + n_x - \operatorname{rank} A_1. \tag{3.8}$$

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

4. Existence and uniqueness of solutions. In this section we prove results on the existence and uniqueness of solutions to (1.5). We aim at establishing the following theorem.

THEOREM 4.1.

(i) *There exists a solution to* (1.5) *if and only if*

$$\boldsymbol{a}_0 \in \operatorname{range}(A_1). \tag{4.1}$$

(ii) *If d is the degree of a solution to* (1.5), *then $D_{\min} \le d \le D_{\max}$.*

(iii) *When* (4.1) *holds, there is a unique solution (p^*, q^*) of minimal degree $D_{\min}$. For this solution, $x_j \neq y_i$ for all indices i, j representing nonzero branch values. Moreover,* $\operatorname{Deg}(q^*) \le n_y - n_x + \operatorname{rank} A_1$ *with equality if $D_{\min} < \operatorname{rank} A_1$.*

(iv) *When* (4.1) *holds, a polynomial pair (p, q) is a solution if and only if $p = p^* r$ and $q = q^* r$, where $r(z)$ is a polynomial satisfying $r(0) = 1$ and* $\operatorname{Deg}(r) \le D_{\max} - D_{\min}$.

(v) *The minimal degree solution is the only solution to* (1.5) *if and only if the matrix A_1 is nonsingular.*

(vi) *Let $\{x_j\}$ and $\{y_j\}$ be a solution to* (1.5). *Then the higher moments defined in* (2.4) *are well defined.*

Let us proceed with several remarks.

REMARK 3. *In particular, it follows from* (i) *that there exists a solution as soon as the matrix A_1 is nonsingular.*

REMARK 4. *Since* (1.5) *is a system of polynomial equations of degree K, one could expect there to be a finite number of solutions, typically K solutions. However, because of the special structure of the equations there is either one unique solution (when A_1 is nonsingular) or infinitely many solutions (when A_1 is singular).*

REMARK 5. *The form $(p^* r, q^* r)$ of solutions can also be stated as follows: All solutions have a core set of values $\{x_j\}$, $j = 1, \dots, \operatorname{Deg}(p^*) = D_{\min}$, and $\{y_i\}$, $i = 1, \dots, \operatorname{Deg}(q^*)$, corresponding to nonzero branch values of the minimal degree solution, where $x_j \neq y_i$ for all those i, j. One can then add an optional set of nonzero branch values $\{x_{D_{\min}+j}\}$ and $\{y_{\operatorname{Deg}(q^*)+j}\}$ for $j = 1, \dots, D_{\max} - D_{\min}$ such that $x_{D_{\min}+j} = y_{\operatorname{Deg}(q^*)+j}$.*

To prove this theorem we first establish some utility results in the next subsection. We then derive different ways of characterizing the solution in section 4.2 which are subsequently used to prove Theorem 4.1 in section 4.3.

4.1. Utility results. We start with a useful lemma on Taylor coefficients for a product of functions.

LEMMA 4.2. *Suppose f, g, and h are analytic functions in a neighborhood of zero satisfying $f(z) = g(z)h(z)$. Let f have the Taylor expansion*

$$f(z) = \sum_{k=0}^{\infty} f_k z^k,$$

and let $\{g_k\}$ and $\{h_k\}$ be the corresponding coefficients for $g(x)$ and $h(x)$, respectively. Then

$$f_k = \sum_{j=0}^{k} g_j h_{k-j}. \tag{4.2}$$

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Proof. Since the functions are analytic, the coefficients are given as

$$f_k = \frac{1}{k!}\frac{d^k}{dz^k}f(z)\Big|_{z=0} = \frac{1}{k!}\frac{d^k}{dz^k}g(z)h(z)\Big|_{z=0} = \frac{1}{k!}\sum_{j=0}^{k} c_{jk}g^{(j)}(0)h^{(k-j)}(0),$$

where $c_{jk} = k!/j!(k-j)!$ are the binomial coefficients. But $g^{(j)}(0) = j!g_j$ and $h^{(k-j)}(0) = (k-j)!h_{k-j}$, and therefore (4.2) follows. □

REMARK 6. *The discrete convolution* (4.2) *is, in fact, precisely an elementwise description of multiplication of a lower triangular $k \times k$ Toeplitz matrix by a vector. In the notation of* [15], *it would read $\boldsymbol{f} = \mathcal{T}(\boldsymbol{g})\boldsymbol{h}$.*

As was already known by Markov, the exponential transform of the moment sequence plays an important role in the analysis of these problems; see, e.g., [1, 2]. We show here that $\{a_k\}$ is a version of the exponential transform of $\{m_k\}$.

LEMMA 4.3. *Suppose $\{m_k\}$ is generated by the polynomials $p(z)$ and $q(z)$ and $\{a_k\}$ is generated by $\{m_k\}$. Let $m(z)$ be defined as*

$$m(z) = m_1 z + \frac{1}{2}m_2 z^2 + \frac{1}{3}m_3 z^3 + \cdots . \tag{4.3}$$

Then, if (3.4) *holds,*

$$e^{m(z)} = \frac{q(z)}{p(z)} = a_0 + a_1 z + a_2 z^2 + \cdots , \tag{4.4}$$

written as its Taylor expansion around $z = 0$.

Proof. Let us first show that $m(z)$ is a well-defined analytic function at zero. We have

$$\begin{aligned} m(z) &= \textstyle\sum_{k=0}^{\infty} \frac{m_k z^k}{k} \\ &= \textstyle\sum_{k=0}^{\infty}\sum_{j=1}^{n_x} \frac{x_j^k z^k}{k} - \sum_{k=0}^{\infty}\sum_{j=1}^{n_y} \frac{y_j^k z^k}{k} \\ &= -\textstyle\sum_{j=1}^{n_x} \log(1 - x_j z) + \sum_{j=1}^{n_y} \log(1 - y_j z). \end{aligned}$$

The last step is allowed when $|z| < 1/\max_{ij}(|x_j|, |y_i|)$, which is true for small enough z since $p(0) \neq 0$. This also shows that the function is analytic at zero. Moreover,

$$e^{m(z)} = \frac{\prod_{j=1}^{n_y}(1 - y_j z)}{\prod_{j=1}^{n_x}(1 - x_j z)} = \frac{q(z)}{p(z)}.$$

Finally, setting $a(z) := \exp(m(z))$ and differentiating gives

$$za'(z) = zm'(z)a(z),$$

where all three functions are analytic at zero. Let $a(z)$ have the Taylor coefficients $\{\tilde{a}_k\}$. Then $za'(z) = \tilde{a}_1 z + 2\tilde{a}_2 z^2 + 3\tilde{a}_3 z^3 \cdots$ and clearly $zm'(z) = m_1 z + m_2 z^2 + \cdots$. By Lemma 4.2, for $k \geq 1$,

$$k\tilde{a}_k = \sum_{j=1}^{k} m_j \tilde{a}_{k-j}.$$

Since $\tilde{a}_0 = q(0)/p(0) = 1$, we see that a_k and $\tilde{a}_k$ satisfy the same nonsingular linear system of equations (2.1), and therefore $a_k = \tilde{a}_k$, showing (4.4). □

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

We now have the following basic characterization of a solution.

LEMMA 4.4. *Suppose* $p(z)$ *and* $q(z)$ *are two polynomials satisfying* (3.3), (3.4). *They form a polynomial solution to* (1.5) *if and only if their quotient has the Taylor expansion around* $z = 0$

$$\frac{q(z)}{p(z)} = a_0 + a_1 z + \cdots + a_K z^K + O\left(z^{K+1}\right), \tag{4.5}$$

where $\{a_k\}$ *is generated by* $\{m_k\}$. *Moreover, if* (p, q) *is a solution, then* $(\bar{p}, \bar{q})$ *is also a solution if and only if the pair satisfies* (3.3), (3.4), *and* $\bar{p}/\bar{q} = p/q$, *where these fractions are defined.*

Proof. Let $\{\tilde{m}_k\}$ be generated by p and q, and suppose (4.5) holds. Then, as in the proof of Lemma 4.3 for $1 \leq k \leq K$,

$$k a_k = \sum_{j=1}^{k} \tilde{m}_j a_{k-j}.$$

Since $\{m_k\}$ satisfy the linear system (2.1), we have after subtraction

$$m_n - \tilde{m}_n = -\sum_{k=1}^{n-1} (m_k - \tilde{m}_k) a_{n-k}, \qquad m_1 = \tilde{m}_1,$$

for $n = 2, \ldots, K$. By induction $\tilde{m}_k = m_k$ for $1 \leq k \leq K$, showing that (p, q) solves (1.5). On the other hand, if (p, q) is a solution, then (4.5) must hold by (4.4) in Lemma 4.3.

For the last statement, the "if" part is obvious since both pairs then satisfy (4.5). To show the "only if" part, suppose both (p, q) and $(\bar{p}, \bar{q})$ are solutions. By definition they satisfy (3.3), (3.4), and by (4.5),

$$\frac{\bar{q}(z)}{\bar{p}(z)} - \frac{q(z)}{p(z)} = \frac{\bar{q}(z)p(z) - \bar{p}(z)q(z)}{\bar{p}(z)p(z)} = O(z^{K+1}).$$

Since $\bar{p}(0)p(0) = 1$, we must have that $(\bar{q}(z)p(z) - \bar{p}(z)q(z))/z^{K+1}$ is bounded as $z \to 0$. But since the degree of $\bar{q}p - \bar{p}q$ is at most $K = n_x + n_y$, this is possible only if it is identically zero. Hence $\bar{q}(z)p(z) = \bar{p}(z)q(z)$, which concludes the proof. ◻

4.2. Characterization of the solution. In this section we show three propositions that characterize solutions to (1.5) in terms of polynomials, coefficient vectors, and the column vectors of the A matrix in (3.6). We start by expressing the uniqueness properties of the solution in terms of its polyomial representation.

PROPOSITION 4.5. *Suppose the pairs* (p, q) *and* $(\bar{p}, \bar{q})$ *are both polynomial solutions to* (1.5). *Then the following hold.*

(i) $\mathrm{Deg}(p) - \mathrm{Deg}(q) = \mathrm{Deg}(\bar{p}) - \mathrm{Deg}(\bar{q})$.

(ii) *If* $\mathrm{Deg}(\bar{p}) \leq \mathrm{Deg}(p)$, *and if there is no polynomial* $r(z)$ *such that* $p = \bar{p}r$, *then there is another solution* $(\tilde{p}, \tilde{q})$ *with* $\mathrm{Deg}(\tilde{p}) < \mathrm{Deg}(p)$. *In particular, if* $\mathrm{Deg}(p) = \mathrm{Deg}(\bar{p})$ *but* $p \neq \bar{p}$, *there is such a lower degree solution.*

(iii) *If* $\mathrm{Deg}(\bar{p}) \leq \mathrm{Deg}(p)$, *any polynomial pair* $(\bar{p}r, \bar{q}r)$ *is a solution if* $r(z)$ *is a polynomial satisfying* $r(0) = 1$ *and* $\mathrm{Deg}(r) \leq \mathrm{Deg}(p) - \mathrm{Deg}(\bar{p})$. *In particular, if* $\mathrm{Deg}(\bar{p}) \leq m \leq \mathrm{Deg}(p)$, *there is a solution* $(\tilde{p}, \tilde{q})$ *with* $\mathrm{Deg}(\tilde{p}) = m$.

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Proof.

(i) The statement follows directly from Lemma 4.4, since $\bar{q}p = \bar{p}q$ implies that

$$\mathrm{Deg}(\bar{q}) + \mathrm{Deg}(p) = \mathrm{Deg}(\bar{p}) + \mathrm{Deg}(q).$$

(ii) We let

$$p(z) = r_p(z)\bar{p}(z) + s_p(z), \qquad q(z) = r_q(z)\bar{q}(z) + s_q(z)$$

be the unique polynomial decomposition of (p, q) such that r_p, r_q, s_p, s_q are polynomials, $\mathrm{Deg}(s_p) < \mathrm{Deg}(\bar{p})$, and $\mathrm{Deg}(s_q) < \mathrm{Deg}(\bar{q})$. Since $\bar{p}q = p\bar{q}$ by Lemma 4.4, we get

$$\bar{p}\bar{q}(r_q - r_p) = \bar{q}s_p - \bar{p}s_q.$$

Unless $r_q = r_p$, the degree of the left-hand side is at least $\mathrm{Deg}(\bar{p}) + \mathrm{Deg}(\bar{q})$, while the degree of the right-hand side is at most

$$\max\left(\mathrm{Deg}(\bar{q}) + \mathrm{Deg}(s_p),\, \mathrm{Deg}(\bar{p}) + \mathrm{Deg}(s_q)\right) < \mathrm{Deg}(\bar{q}) + \mathrm{Deg}(\bar{p}).$$

Hence, $r_q = r_p$ and $\bar{q}s_p = \bar{p}s_q$. Since $\bar{q}, \bar{p} \not\equiv 0$, it follows that s_p and s_q are either both zero or both nonzero. Suppose $s_p \not\equiv 0$ and $s_q \not\equiv 0$. Write $s_p(z) = z^{m_p}\tilde{s}_p(z)$ and $s_q(z) = z^{m_q}\tilde{s}_q(z)$, where $\tilde{s}_p(0) \neq 0$ and $\tilde{s}_q(0) \neq 0$. Since

$$z^{m_p}\tilde{s}_p(z)\bar{q}(z) = z^{m_q}\tilde{s}_q(z)\bar{p}(z)$$

and also $\bar{q}(0) = \bar{p}(0) = 1$, the lowest degree term in the left- and right-hand side polynomials are z^{m_p} and z^{m_q}, respectively, and therefore $m_p = m_q$. Consequently,

$$\tilde{s}_p(z)\bar{q}(z) = \tilde{s}_q(z)\bar{p}(z)$$

and $\tilde{s}_p(0) = \tilde{s}_q(0)$. We can then take $\tilde{p}(z) = \tilde{s}_p(z)/\tilde{s}_p(0)$ and $\tilde{q}(z) = \tilde{s}_q(z)/\tilde{s}_q(0)$. They satisfy

$$\tilde{p}(z)\bar{q}(z) = \tilde{q}(z)\bar{p}(z), \qquad \tilde{p}(0) = \tilde{q}(0) = 1,$$

while $\mathrm{Deg}(\tilde{p}) = \mathrm{Deg}(\tilde{s}_p) \leq \mathrm{Deg}(s_p) < \mathrm{Deg}(p)$ and similarly $\mathrm{Deg}(\tilde{q}) < \mathrm{Deg}(q) \leq n_y$. Hence $(\tilde{p}, \tilde{q})$ is a polynomial solution by Lemma 4.4. It has degree strictly less than (p, q), which shows the first statement in (ii). If $\mathrm{Deg}(p) = \mathrm{Deg}(\bar{p})$ and $p \neq \bar{p}$, then there is no $r(z)$ satisfying the requirements, showing the second statement in (ii).

(iii) We finally let $r(z)$ be any polynomial with $\mathrm{Deg}(r) \leq \mathrm{Deg}(p) - \mathrm{Deg}(\bar{p})$ and $r(0) = 1$. We then set $\tilde{p} = \bar{p}r$ and $\tilde{q} = \bar{q}r$. These polynomials trivially satisfy (3.4) and (4.5). Since $\mathrm{Deg}(\tilde{p}) = \mathrm{Deg}(r) + \mathrm{Deg}(\bar{p}) \leq \mathrm{Deg}(p) \leq n_x$ and

$$\mathrm{Deg}(\tilde{q}) = \mathrm{Deg}(r) + \mathrm{Deg}(\bar{q}) \leq \mathrm{Deg}(p) - \mathrm{Deg}(\bar{p}) + \mathrm{Deg}(\bar{q}) = \mathrm{Deg}(q) \leq n_y,$$

they also satisfy (3.3) and thus are a polynomial solution by Lemma 4.4. In particular, we can take $r(z)$ of degree m. ∎

A solution to (1.5) can also be characterized in terms of the coefficient vectors. We have the following proposition.

PROPOSITION 4.6. *The pair* $\boldsymbol{c} = (c_0, \ldots, c_{n_x})^T \in \mathbb{R}^{n_x+1}$ *and* $\boldsymbol{d} = (d_0, \ldots, d_{n_y})^T \in \mathbb{R}^{n_y+1}$ *is a coefficient solution to* (1.5) *if and only if*

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

(i) $c_0 = 1$,
(ii) *$\boldsymbol{c}$ is in the null-space of A, and*
(iii)

$$d_k = \sum_{j=0}^{\min(k,n_x)} c_j a_{k-j}, \qquad k = 0, \dots, n_y. \tag{4.6}$$

Proof. Suppose first that $\boldsymbol{c}$ is in the null-space of A, $c_0 = 1$, and $\{d_k\}$ is given by (4.6). Extend the coefficient sequences by setting $c_k = 0$ for $k > n_x$ and $d_k = 0$ for $k > n_y$. Since $\boldsymbol{c}$ is in the null-space of A, we get $\sum_{j=0}^{k} c_j a_{k-j} = 0$ when $n_y + 1 \le k \le n_x + n_x = K$, and in conclusion

$$d_k = \sum_{j=0}^{k} c_j a_{k-j}, \qquad k = 0, \dots, K. \tag{4.7}$$

Upon noting that $\{c_k\}_{k=0}^{\infty}$ and $\{d_k\}_{k=0}^{\infty}$ are the Taylor coefficients of P_c and P_d, and since $P_c(0) = c_0 = 1$, $P_d(0) = d_0 = a_0 c_0 = 1$, Lemma 4.2 shows that

$$P_d(z) = P_c(z)\left[a_0 + a_1 z + \cdots + a_K z^K + O\left(z^{K+1}\right)\right], \tag{4.8}$$

and by Lemma 4.4 we have that (P_c, P_d) is a solution to (1.5). Conversely, if (P_c, P_d) is a solution, then $c_0 = P_c(0) = 1$, and by Lemma 4.2 we get that (4.7) holds. For $k = n_y + 1, \dots, K$ this also implies that $\boldsymbol{c}$ is in the null-space of A. ∎

The final proposition of this section relates the degree of the solution to the column vectors of A and the linear spaces they span.

PROPOSITION 4.7. *Let $V_j = \operatorname{span}\{\boldsymbol{a}_1, \dots, \boldsymbol{a}_j\}$ and $V_j^0 = \operatorname{span}\{\boldsymbol{a}_0, \dots, \boldsymbol{a}_j\}$. Set $V_0 = V_{-1}^0 = \emptyset$. Then the following hold.*

(i) *There is a solution if and only if $\boldsymbol{a}_0 \in V_{n_x} = \mathrm{Range}(A_1)$.*
(ii) *There is a solution of degree $j \ge 0$ if and only if*

$$\boldsymbol{a}_0 \in V_j \quad \text{and} \quad \boldsymbol{a}_j \in V_{j-1}^0. \tag{4.9}$$

(iii) *When $\boldsymbol{a}_0 \in V_{n_x}$, then*

$$\boldsymbol{a}_0 \in V_d, \qquad V_d^0 = V_d$$

if and only if $d \ge D_{\max}$.

(iv) *When $\boldsymbol{a}_0 \in V_{n_x}$, the vectors*

$$\boldsymbol{a}_1, \dots, \boldsymbol{a}_{D_{\min}}$$

(when $D_{\min} > 0$),

$$\boldsymbol{a}_{D_{\max}+1}, \dots, \boldsymbol{a}_{n_x}$$

(when $D_{\max} < n_x$) are all linearly independent. Moreover,

$$\boldsymbol{a}_j \in V_{D_{\min}}, \quad V_j = V_{D_{\min}}, \qquad j = D_{\min}, \dots, D_{\max}.$$

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Proof.

(i) By Proposition 4.6 there exists a solution to (1.5) if and only if there is a coefficient vector $\boldsymbol{c} = (1, \boldsymbol{c}')^T$ in the null-space of A, i.e.,

$$A\boldsymbol{c} = A_1\bar{\boldsymbol{c}} + \boldsymbol{a}_0 = 0.$$

But such a vector $\bar{\boldsymbol{c}}$ exists if and only if $\boldsymbol{a}_0$ is in the range of A_1. This shows (i).

(ii) Again by Proposition 4.6 there is a solution of degree j if and only if there is a vector $\boldsymbol{c} = (c_0, c_1, \ldots, c_j, 0, \ldots, 0)^T$ such that

$$0 = A\boldsymbol{c} = c_0\boldsymbol{a}_0 + c_1\boldsymbol{a}_1 + \cdots + c_j\boldsymbol{a}_j, \tag{4.10}$$

with $c_j \neq 0$ and $c_0 = 1$. For $j = 0$ this is clearly equivalent to $\boldsymbol{a}_0 = 0$ or $\boldsymbol{a}_0 \in V_0 = V_{-1}^0$. For $j > 0$ the existence of c_j-coefficients satisfying (4.10) is equivalent to the left condition in (4.9). Moreover, if $\boldsymbol{a}_j \neq V_{j-1}^0 = \text{span}\{\boldsymbol{a}_0, \ldots, \boldsymbol{a}_{j-1}\}$, then we must have $c_j = 0$ to satisfy (4.10), and $\boldsymbol{c}$ cannot represent a solution of degree j. On the other hand, if $c_j = 0$ and $\boldsymbol{a}_j = c_0'\boldsymbol{a}_0 + \cdots + c_{j-1}'\boldsymbol{a}_{j-1}$ for some nonzero coefficients c_k', then $\boldsymbol{a}_0 + c_1''\boldsymbol{a}_1 + \cdots + c_{j-1}''\boldsymbol{a}_{j-1} + \boldsymbol{a}_j = 0$, with $c_k'' = (1 + c_0')c_k - c_k'$, represents a solution of degree j. This shows (ii).

(iii) The statement is obvious in case $D_{\min} = 0$. If $D_{\min} > 0$, there are scalars such that

$$\boldsymbol{a}_0 = v_1\boldsymbol{a}_1 + \cdots + v_{D_{\min}}\boldsymbol{a}_{D_{\min}}, \tag{4.11}$$

by (3.7). Hence, $\boldsymbol{a}_0 \in V_{D_{\min}}$ and since the V_j spaces are nested, $V_j \subset V_{j+1}$, we have $\boldsymbol{a}_0 \in V_d$ for $d \geq D_{\min}$. Moreover, the minimal property of $D_{\min}$ ensures that $v_{D_{\min}} \neq 0$ in (4.11), so that $\boldsymbol{a}_0 \notin V_d$ when $d < D_{\min}$.

(iv) To show that when $D_{\min} > 0$ the vectors $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_{D_{\min}}$ are linearly independent, we use (4.11) and note that $P_c(z)$ with $\boldsymbol{c} = (1, -v_1, \ldots, -v_{D_{\min}}, 0, \ldots, 0)^T$ is a polynomial solution to (1.5). Suppose now that there are nonzero coefficients c_j' such that

$$c_1'\boldsymbol{a}_1 + \cdots + c_{D_{\min}}'\boldsymbol{a}_{D_{\min}} = 0.$$

Then $P_{c'}$ with $\boldsymbol{c}' = (1, c_1' - v_1, \ldots, c_{D_{\min}} - v_{D_{\min}}, 0, \ldots, 0)^T$ is another polynomial solution to (1.5). Moreover, by the minimality property of $D_{\min}$ we must have $c_{D_{\min}} - v_{D_{\min}} \neq 0$ and therefore $\text{Deg}(P_c) = \text{Deg}(P_{c'}) = D_{\min}$. But by (ii) in Proposition 4.5 this implies that there is yet another solution $P_{c''}$ of degree strictly less than $D_{\min}$. Hence, there are coefficients c_j'' such that

$$\boldsymbol{a}_0 + c_1''\boldsymbol{a}_1 + \cdots + c_d''\boldsymbol{a}_d = 0,$$

with $d < D_{\min}$, contradicting (3.7). The vectors must therefore be linearly independent.

Suppose $D^* \geq D_{\min}$ is the highest degree of an existing solution. Since $P_c(z)$ is a solution of degree $D_{\min}$, we get from (iii) in Proposition 4.5 that there are solutions of all intermediate degrees $D_{\min}, \ldots, D^*$. Hence, from (ii), $\boldsymbol{a}_j \in V_{j-1}^0$ for $j = D_{\min}, \ldots, D^*$, and, from (iii), $\boldsymbol{a}_j \in V_{j-1}$ for $j = D_{\min}+1, \ldots, D^*$. Noting that if $\boldsymbol{a}_{j+1} \in V_j$, then $V_j = V_{j+1}$, we can conclude inductively that $V_{D_{\min}} = \cdots = V_{D^*}$ and $\boldsymbol{a}_j \in V_{D_{\min}}$ for $j = D_{\min}, \ldots, D^*$. We now have three different cases.

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

1. If $D^* = n_x$, then $V_{D_{\min}} = V_{n_x}$, and by (3.8) we get $D^* = \operatorname{rank} A_1 - D_{\min} + D_{\max} = \dim V_{n_x} - D_{\min} + D_{\max} = \dim V_{D_{\min}} - D_{\min} + D_{\max} = D_{\max}$ since either $\boldsymbol{a}_1, \dots, \boldsymbol{a}_{D_{\min}}$ are linearly independent or $D_{\min} = 0$ and $V_{D_{\min}} = \emptyset$. This shows (iv) for $D^* = n_x$.
2. If $D^* < n_x$ and $D_{\min} = 0$, then $V_{D_{\min}} = V_{D^*} = \emptyset$, and

   $$V_{n_x} = \operatorname{span}\{\boldsymbol{a}_{D^*+1}, \dots, \boldsymbol{a}_{n_x}\}. \tag{4.12}$$

   Suppose there are nonzero coefficients α_k such that

   $$\alpha_{D^*+1}\boldsymbol{a}_{D^*+1} + \cdots + \alpha_{n_x}\boldsymbol{a}_{n_x} = 0,$$

   and let k^* be the highest index of all nonzero coefficients, $\alpha_{k^*} \neq 0$. Then $\boldsymbol{a}_{k^*} \in V^0_{k^*-1}$ and there is a solution of degree k^* by (ii), which is in contradiction to the definition of D^*. Hence, the vectors in (4.12) must be linearly independent and

   $$D^* = n_x - \dim V_{n_x} = D_{\min} + n_x - \operatorname{rank} A_1 = D_{\max},$$

   showing (iv) for this case.
3. If $D^* < n_x$ and $D_{\min} > 0$, we have

   $$V_{n_x} = \operatorname{span}\{\boldsymbol{a}_1, \dots, \boldsymbol{a}_{D_{\min}}, \boldsymbol{a}_{D^*+1}, \dots, \boldsymbol{a}_{n_x}\}. \tag{4.13}$$

   Suppose there are nonzero coefficients α_k such that

   $$\alpha_1\boldsymbol{a}_1 + \cdots + \alpha_{D_{\min}}\boldsymbol{a}_{D_{\min}} + \cdots + \alpha_{D^*+1}\boldsymbol{a}_{D^*+1} + \cdots + \alpha_{n_x}\boldsymbol{a}_{n_x} = 0.$$

   Since $\boldsymbol{a}_1, \dots, \boldsymbol{a}_{D_{\min}}$ are linearly independent, at least one α_k with $k > D^*$ must be nonzero. By the same argument as above in case 2 we then get a contradiction, and the vectors in (4.13) must be linearly independent. Hence,

   $$D^* = D_{\min} + n_x - \dim V_{n_x} = D_{\min} + n_x - \operatorname{rank} A_1 = D_{\max},$$

   showing this final case. ∎

4.3. Proof of Theorem 4.1. To prove Theorem 4.1, we essentially have to combine the results from Propositions 4.5 and 4.7. The statement (i) is given directly by (i) in the latter. For the remaining points we have the following.

(ii) From (ii) in Proposition 4.7 we see that $\boldsymbol{a}_0 \in V_d$ and $\boldsymbol{a}_d \in V^0_{d-1}$. It follows from (iii) in Proposition 4.7 that $d \geq D_{\min}$. On the other hand, if $D_{\max} < n_x$ and $d > D_{\max}$, it says that $V^0_{d-1} = V_{d-1}$. Hence, $\boldsymbol{a}_d \in V_{d-1}$, which contradicts the linear independence of $\boldsymbol{a}_{D_{\max}}, \dots, \boldsymbol{a}_{n_x}$ established in point (iv) of Proposition 4.7.

(iii) We note that by (3.7) there are scalars $v_1, \dots, v_{D_{\min}}$ such that

$$\boldsymbol{a}_0 = v_1\boldsymbol{a}_1 + \cdots + v_{D_{\min}}\boldsymbol{a}_{D_{\min}}. \tag{4.14}$$

Hence, $\boldsymbol{a}_0 \in V_{D_{\min}}$, and since $v_{D_{\min}} \neq 0$, we also have $\boldsymbol{a}_{D_{\min}} \in V^0_{D_{\min}}$. By (ii) in Proposition 4.7 there is thus a solution of degree $D_{\min}$ which we denote (p^*, q^*). Since $\boldsymbol{a}_1, \dots, \boldsymbol{a}_{D_{\min}}$ are linearly independent by (iii) in Proposition 4.7, the coefficients in (4.14) are unique, and therefore the $D_{\min}$-degree

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

solution is also unique. Moreover, suppose that $x_j = y_i = x^* \neq 0$ for some i, j. Then p^* and q^* would have a common factor $(1-zx^*)$, and by Lemma 4.4 $\bar{p}(z) := p^*(z)/(1-zx^*)$ and $\bar{q}(z) := q^*(z)/(1-zx^*)$ would also be a solution. But this is impossible since $\operatorname{Deg}(\bar{p}) < \operatorname{Deg}(p^*) = D_{\min}$. By (iv), shown below, a solution is given by (p^*r, q^*r), where $r(0) = 1$ and $\operatorname{Deg}(r) = D_{\max} - D_{\min}$. Hence $n_y \geq \operatorname{Deg}(q^*r) = \operatorname{Deg}(q^*) + n_x - \operatorname{rank} A_1$. Suppose finally that $D_{\min} < \operatorname{rank} A_1$ and that $\operatorname{Deg}(q^*) < n_y - n_x + \operatorname{rank} A_1$. Let $\operatorname{Deg}(r) = D_{\max} + 1 - D_{\min}$. Then (p^*r, q^*r) is still a solution by Lemma 4.4 since (p^*, q^*) is a solution, $\operatorname{Deg}(p^*r) = D_{\max} + 1 = n_x + D_{\min} + 1 - \operatorname{rank} A_1 \leq n_x$, and

$$\operatorname{Deg}(q^*r) < n_y - n_x + \operatorname{rank} A_1 + D_{\max} + 1 - D_{\min} = n_y + 1.$$

This contradicts (ii) and therefore $\operatorname{Deg}(q^*) = n_y - n_x + \operatorname{rank} A_1$, concluding the proof of (iii).

(iv) We first note that there exists a solution of degree $D_{\max}$ by Proposition 4.7 since, if $D_{\max} > D_{\min}$, we have $\boldsymbol{a}_0 \in V^0_{D_{\max}-1}$ and $\boldsymbol{a}_{D_{\max}} \in V_{D_{\min}} = V_{D_{\max}-1} = V^0_{D_{\max}-1}$. Hence, (iii) in Proposition 4.5 shows that any polynomial pair of the stated type is a solution. On the other hand, if the polynomial solution is not of this type, then (ii) in Proposition 4.5 says there is a solution of degree strictly less than $D_{\min}$, contradicting (ii) above.

(v) We suppose first that A_1 is nonsingular. Then $\operatorname{rank} A_1 = n_x$ so that $D_{\min} = D_{\max}$ and the uniqueness is given by (iii) above. If, on the contrary, A_1 is singular, then $D_{\max} > D_{\min}$, and since we can then pick infinitely many polynomials $r(z)$ in (iv), we have infinitely many solutions.

(vi) This is a consequence of (iv). The solution can be represented by (p^*r, q^*r) for some polynomial $r(z)$ with $r(0) = 1$. Let $1/x_j$ for $j = 1, \dots, D_{\min}$ and $1/y_j$ for $j = 1, \dots, \operatorname{Deg}(q^*)$ be the roots of $p^*(z)$ and $q^*(z)$, respectively. Let $1/z_j$ for $j = 1, \dots, \operatorname{Deg}(r)$ be the roots of $r(z)$. Then

$$m_k = \sum_{j=1}^{D_{\min}} x_j^k + \sum_{j=1}^{\operatorname{Deg}(r)} z_j^k - \sum_{j=1}^{\operatorname{Deg}(q^*)} y_j^k - \sum_{j=1}^{\operatorname{Deg}(r)} z_j^k = \sum_{j=1}^{D_{\min}} x_j^k - \sum_{j=1}^{\operatorname{Deg}(q^*)} y_j^k,$$

which is independent of $r(z)$ and uniquely determined because (p^*, q^*) is unique.

5. Proof of Theorem 2.1. We can now use the results in section 4 to prove Theorem 2.1.

(i)–(ii) To show the statements about Algorithms 1 and 2 we consider the reduced problem

$$m_k = \sum_{j=1}^{\tilde{n}_x} \tilde{x}_j^k - \sum_{j=1}^{\tilde{n}_y} \tilde{y}_j^k, \qquad k = 1, \dots, \tilde{K}, \tag{5.1}$$

where $\tilde{n}_x = \operatorname{rank} A_1 \leq n_x$, $\tilde{n}_y = n_y - n_x + \tilde{n}_x \leq n_y$, and $\tilde{K} = \tilde{n}_x + \tilde{n}_y \leq K$. The moments m_k in the left-hand side are the same as in (1.5). First, we consider the minimal solution (p^*, q^*) of (1.5). By (iv) in Proposition 4.7 we must have $\operatorname{Deg}(p^*) = D_{\min} \leq \operatorname{rank} A_1 = \tilde{n}_x$. Moreover, by (iii) in Theorem 4.1,

$$\operatorname{Deg}(q^*) \leq n_y - n_x + \operatorname{rank} A_1 = \tilde{n}_y.$$

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

It follows from Lemma 4.4 that (p^*, q^*) is also a solution to (5.1). Second, let $(\tilde{p}^*, \tilde{q}^*)$ be the minimal degree solution to (5.1). Then by (iv) in Theorem 4.1 there is a polynomial $r(z)$ with $r(0) = 1$ such that $p^* = \tilde{p}^* r$ and $q^* = \tilde{q}^* r$. But then $(\tilde{p}^*, \tilde{q}^*)$ is also a solution to (1.5) by Lemma 4.4. By the uniqueness of the minimal degree solution of (1.5) it follows that $r \equiv 1$ and $p^* = \tilde{p}^*$, $q^* = \tilde{q}^*$. Suppose now that there is another polynomial $r(z)$ with $r(0) = 1$, $\mathrm{Deg}(r) > 0$ such that $(p^* r, q^* r)$ is a solution to (5.1). Then $\mathrm{Deg}(p^* r) = D_{\min} + \mathrm{Deg}(r) \leq \tilde{n}_x = \operatorname{rank} A_1$. Hence, $D_{\min} < \operatorname{rank} A_1$, and therefore by (iii) in Theorem 4.1 we have $\mathrm{Deg}(q^*) = n_y - n_x + \operatorname{rank} A_1 = \tilde{n}_y$. Thus, $\mathrm{Deg}(q^* r) > \tilde{n}_y$, which is impossible if $(p^* r, q^* r)$ is a solution. Hence, (p^*, q^*) is the unique solution to (5.1), and therefore $\tilde{A}_1$ is nonsingular by (v) in Theorem 4.1.

Since $\tilde{A}_1$ is invertible, the generalized eigenvalue problem (2.2) and $\boldsymbol{c}'$ are well defined. Moreover, we can construct $\tilde{A}_1^{-1} \tilde{A}_0$. By (2.3),

$$\tilde{A}_1^{-1} \tilde{A}_0 = \begin{pmatrix} -c_1 & 1 & 0 & \cdots & 0 \\ -c_2 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ -c_{\tilde{n}_x - 1} & 0 & 0 & \ddots & 1 \\ -c_{\tilde{n}_x} & 0 & 0 & \cdots & 0 \end{pmatrix},$$

which is a companion matrix. It is well known that for those matrices the elements in the first column are the coefficients of its characteristic polynomial. This is shown as follows: let M_{ij} be the minor of $V := zI - \tilde{A}_1^{-1} \tilde{A}_0$, i.e., the determinant of the matrix obtained by removing row i and column j. Then, the determinant can be expanded by minors, for any j:

$$\det(V) = \sum_{i=1}^{\tilde{n}_x} (-1)^{i+j} v_{ij} M_{ij}, \qquad V = \{v_{ij}\}.$$

Taking $j = 1$, we get $M_{i,1} = \det(\operatorname{diag}(z, \dots, z, -1, \dots, -1))$ with $i - 1$ occurrences of -1, so that $M_{i,1} = z^{\tilde{n}_x - i} (-1)^{i-1}$. Therefore,

$$\begin{aligned} \det(V) &= (-1)^2 (c_1 + z) M_{1,1} + \textstyle\sum_{i=2}^{\tilde{n}_x} (-1)^{i+1} c_i M_{i,1} \\ &= c_1 z^{\tilde{n}_x - 1} + z^{\tilde{n}_x} + \textstyle\sum_{i=2}^{\tilde{n}_x} c_i z^{\tilde{n}_x - i} \\ &= P(z), \end{aligned}$$

which is exactly (2.6). This shows that the results of Algorithms 1 and 2 are identical, since the generalized eigenvalues in (2.2) are exactly the roots of $P(z)$.

It remains to show what the roots are. Let $\tilde{A} = [\tilde{a}_0 \ \tilde{A}_1]$ be the A matrix related to (5.1). Clearly, $\boldsymbol{c} = (1, \boldsymbol{c}'^T)^T$ is in the null-space of $\tilde{A}$, and hence $P_c(z)$ is the unique solution to (5.1). But for $z \neq 0$,

$$\begin{aligned} P(z) &= c_{\tilde{n}_x} + c_{\tilde{n}_x - 1} z + \cdots + c_1 z^{\tilde{n}_x - 1} + z^{\tilde{n}_x} \\ &= z^{\tilde{n}_x} \left(\frac{c_{\tilde{n}_x}}{z^{\tilde{n}_x}} + \frac{c_{\tilde{n}_x - 1}}{z^{\tilde{n}_x - 1}} + \cdots + \frac{c_1}{z} + 1 \right) \\ &= z^{\tilde{n}_x} P_c(1/z) \\ &= z^{\tilde{n}_x} (1 - x_1/z)(1 - x_2/z) \cdots (1 - x_{D_{\min}}/z) \\ &= z^{\tilde{n}_x - D_{\min}} (z - x_1)(z - x_2) \cdots (z - x_{D_{\min}}), \end{aligned}$$

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

which extends to $z = 0$ by continuity. This concludes the proof of points (i) and (ii).

(iii) Let (p, q) be a polynomial solution to (1.5) and $\boldsymbol{c}$ the corresponding coefficient solution. From Lemma 4.3 we have

$$q(z) = p(z)e^{m(z)},$$

where $m(z)$ is defined in (4.3). For the $(K+1)$th Taylor coefficient of the left- and right-hand sides we have by Lemmas 4.3 and 4.2

$$0 = \sum_{j=0}^{n_x} c_j a_{K+1-j} \quad \Rightarrow \quad a_{K+1} = -\sum_{j=1}^{n_x} a_{K+1-j} c_j, \tag{5.2}$$

since the kth Taylor coefficient of q and p is zero for $k > n_x$ and $k > n_y$, respectively. Finally, the last row of (2.1) extended to size $K + 1$ gives

$$m_{K+1} = (K+1)a_{K+1} - \sum_{j=1}^{K} m_j a_{K+1-j}.$$

Together the last two equations show point (iii).

6. Properties of $\boldsymbol{A_1}$ and Markov's theorem. We now look in more detail at the structure of the A_1 matrix. In particular, we look at the implications of A_1R being positive definite. Then we get an explicit simplified formula for the matrix, and our results also shed some light on the relationship of our results to the classical Markov theorem on the existence and uniqueness of solutions to the finite moment problem (1.1) discussed in the introduction. For this we need to define the matrix

$$R = \begin{pmatrix} & & 1 \\ & \cdots & \\ 1 & & \end{pmatrix}$$

and note that left (right) multiplication by R reverses the order of rows (columns) of a matrix. In our notation we can then formulate Markov's theorem as follows.

THEOREM 6.1 (Markov). *Suppose $K = 2n$ is even and $n = n_x = n_y$. There is a unique piecewise continuous function $f(x)$ satisfying*

$$m_k = k \int_{\mathbb{R}} x^{k-1} f(x)dx, \qquad 0 \le f \le 1, \qquad k = 1, \dots, K, \tag{6.1}$$

if A_1R is symmetric positive definite and the matrix

$$\begin{pmatrix} \boldsymbol{a}_0 & A_1 \\ a_{K+1} & \boldsymbol{a}_0^T \end{pmatrix} \tag{6.2}$$

is singular. This f is of the form in (1.2), (1.3).

REMARK 7. *The theorem does not rule out other forms of $f(x)$ a priori, and without the second condition in* (6.2) *such solutions are indeed possible. It considers only the case $n_x = n_y$, i.e., problem* (1.4), *and says nothing about the possibility of other solution types, e.g., when the $\{x_j\}$ and $\{y_j\}$ are not interlaced as in* (1.3).

We start by introducing some new notation that will be used throughout this section. If $\{x_j\}$ and $\{y_j\}$ are a solution of (1.5) and (p, q) is the corresponding

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

polynomial solution as defined in (3.1), (3.2), we can introduce the new polynomials $p_r(z) = z^{n_x} p(1/z)$ and $q_r(z) = z^{n_y} q(1/z)$ to describe the solution. Defining them by continuity at $z = 0$, we have

$$p_r(z) = (z - x_1)\cdots(z - x_{n_x}), \qquad q_r(z) = (z - y_1)\cdots(z - y_{n_y}). \tag{6.3}$$

Furthermore, we assume that the number of *distinct* roots of p_r (x_j-branch values) is $\tilde{n}$. We also order the roots such that we can write

$$p_r(z) = (z - x_1)^{1+\eta_1}(z - x_2)^{1+\eta_2}\cdots(z - x_{\tilde{n}})^{1+\eta_{\tilde{n}}},$$

where $1 + \eta_j$ is the multiplicity of the root x_j, so that

$$n_x = \mathrm{Deg}(p_r) = \tilde{n} + \sum_{\ell=1}^{\tilde{n}} \eta_\ell.$$

We start the analysis with a lemma giving explicit expressions for the a_k-values.

LEMMA 6.2. *For $k \geq 0$,*

$$a_{n_y-n_x+1+k} = \sum_{j=1}^{\tilde{n}} \frac{1}{\eta!} \lim_{z\to x_j} \frac{d^{\eta_j}}{dz^{\eta_j}} \frac{(z - x_j)^{1+\eta_j} z^k q_r(z)}{p_r(z)}. \tag{6.4}$$

Proof. This result follows from an application of the residue theorem in complex analysis as follows. Let C_r be the circle in the complex plane with radius r. Since the roots of $p(z)$ are nonzero, the function q/p is analytic within and on C_ε if ε is taken small enough, and the Cauchy integral formula gives

$$a_k = \begin{cases} \frac{1}{k!}\frac{d^k}{dz^k}\frac{q(z)}{p(z)}\Big|_{z=0}, & k \geq 0, \\ 0, & k < 0, \end{cases} = \frac{1}{2\pi i}\oint_{C_\varepsilon} \frac{q(z)}{p(z)z^{k+1}}dz.$$

Setting

$$f(z) := \frac{q_r(z)}{p_r(z)} = \frac{z^{n_y-n_x} q(1/z)}{p(1/z)} \tag{6.5}$$

and changing variable $z \to 1/z$, we get

$$a_{n_y-n_x+1+k} = \frac{1}{2\pi i}\oint_{C_\varepsilon} \frac{q(z)}{p(z)z^{n_y-n_x+k+2}}dz = \frac{1}{2\pi i}\oint_{C_\varepsilon} \frac{f(1/z)}{z^{k+2}}dz = \frac{1}{2\pi i}\oint_{C_{1/\varepsilon}} z^k f(z)dz.$$

Hence, $a_{n_y-n_x+1+k}$ is given by the sum of the residues of $z^k f(z)$ (assuming we take small enough ε). By (6.5) and the restriction $k \geq 0$ we see that its poles are located at the x_j-values and they have multiplicities $1+\eta_j$ at x_j. Then (6.4) follows from the residue formula for a pole of a function $g(z)$ at z^* with multiplicity $\eta + 1$,

$$\mathrm{Res}(g, z^*) = \frac{1}{\eta!}\lim_{z\to z^*}\frac{d^\eta}{dz^\eta}(z - z^*)^{1+\eta} g(z). \qquad \square$$

When the branch values $\{x_j\}$ are *distinct*, the expression for the a_k elements simplifies. They can then be expressed as sums of the powers of $\{x_j\}$ in a way similar to the moments m_k, but with weights different from one. We can also give a more

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

concise description of the matrices A_0 and A_1, which can be factorized into a product of Vandermonde and diagonal matrices. More precisely, we let V be the Vandermonde matrix

$$V = \begin{pmatrix} 1 & 1 & \cdots & 1 \\ x_1 & x_2 & \cdots & x_{n_x} \\ x_1^2 & x_2^2 & \cdots & x_{n_x}^2 \\ \vdots & \cdots & \cdots & \vdots \\ x_1^{n_x-1} & x_2^{n_x-1} & \cdots & x_{n_x}^{n_x-1} \end{pmatrix}$$

and introduce the diagonal matrices

$$W = \begin{pmatrix} w_1 & & \\ & \ddots & \\ & & w_{n_x} \end{pmatrix}, \qquad X = \begin{pmatrix} x_1 & & \\ & \ddots & \\ & & x_{n_x} \end{pmatrix},$$

where w_j are the weights defined as

$$w_j = \frac{q_r(x_j)}{p_r'(x_j)}. \tag{6.6}$$

(Note that p_r has only simple roots when $\{x_j\}$ are distinct, so $p_r'(x_j) \neq 0$.) Then we can show the following.

PROPOSITION 6.3. *If $\{x_j\}$ are distinct, then for $k \geq 0$,*

$$a_{n_y-n_x+1+k} = \sum_{j=1}^{n_x} w_j x_j^k \tag{6.7}$$

and

$$A_1 R = VWV^T, \qquad A_0 R = VWXV^T. \tag{6.8}$$

Proof. When $\{x_j\}$ are distinct, $\eta_j = 0$ for all j and the expression (6.4) for the x_j-residue simplifies to

$$\lim_{z \to x_j} \frac{(z - x_j) z^k q_r(z)}{p_r(z)} = \frac{x_j^k q_r(x_j)}{p_r'(x_j)}.$$

This shows (6.7). For (6.8) we set $b_k = a_{n_y-n_x+1+k}$. Then

$$A_{1-r} R = \begin{pmatrix} b_r & b_{r+1} & \cdots & b_{r+n_x} \\ b_{r+1} & b_{r+2} & \cdots & b_{r+n_x+1} \\ \vdots & \vdots & \ddots & \vdots \\ b_{r+n_x} & b_{r+n_x+1} & \cdots & b_{r+2n_x} \end{pmatrix} \in \mathbb{R}^{n_x \times n_x}, \qquad r = 0, 1.$$

From (6.7) we then have, for $k \geq 0$,

$$\begin{pmatrix} b_k \\ b_{k+1} \\ \vdots \\ b_{k+n_x} \end{pmatrix} = \sum_{j=1}^{n_x} w_j \begin{pmatrix} x_j^k \\ x_j^{k+1} \\ \vdots \\ x_j^{k+n_x} \end{pmatrix} = \sum_{j=1}^{n_x} w_j x_j^k \begin{pmatrix} 1 \\ x_j \\ \vdots \\ x_j^{n_x} \end{pmatrix} = V \begin{pmatrix} w_1 x_1^k \\ w_2 x_2^k \\ \vdots \\ w_{n_x} x_{n_x}^k \end{pmatrix} = VW \begin{pmatrix} x_1^k \\ x_2^k \\ \vdots \\ x_{n_x}^k \end{pmatrix}.$$

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Consequently,

$$A_{1-r}R = VW \begin{pmatrix} x_1^r & x_1^{r+1} & \dots & x_1^{r+n_x} \\ x_2^r & x_2^{r+1} & \dots & x_2^{r+n_x} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n_x}^r & x_{n_x}^{r+1} & \dots & x_{n_x}^{r+n_x} \end{pmatrix} = VWX^rV^T,$$

which concludes the proof. ∎

We now consider the implications of a positive definite A_1R. It turns out that this is a necessary and sufficient condition to guarantee both distinct $\{x_j\}$-values and positive weights. We get the following theorem.

THEOREM 6.4. *The matrix A_1R is symmetric positive definite if and only if $\{x_j\}$ are distinct and the weights are strictly positive, $w_j > 0$ for $j = 1, \dots, n_x$.*

Proof. We use the same notation as in Lemma 6.2 and set

$$S_j(z) = \frac{1}{\eta_j!}(z - x_j)^{1+\eta_j}\frac{q_r(z)}{p_r(z)}.$$

We note that $S_j(z)$ is smooth and regular close to $z = x_j$. Then, by Lemma 6.2, for $k \geq 0$,

$$a_{n_y-n_x+1+k} = \sum_{j=1}^{\tilde{n}} \lim_{z\to x_j} \frac{d^{\eta_j}}{dz^{\eta_j}} z^k S_j(z).$$

Next, we let $\boldsymbol{v} = (v_1, \dots, v_{n_x})^T$ be an arbitrary vector in $\mathbb{R}^{n_x}$ and recall that $P_v(z)$ is the corresponding $n_x - 1$ degree polynomial

$$P_v(z) = v_1 + v_2 z + \dots + v_{n_x} z^{n_x-1}.$$

Then

$$\boldsymbol{v}^T A_1 R \boldsymbol{v} = \sum_{j=1}^{n_x}\sum_{k=1}^{n_x} v_j v_k a_{n_y-n_x+j+k-1} = \sum_{j=1}^{n_x}\sum_{k=1}^{n_x}\sum_{\ell=1}^{\tilde{n}} \lim_{z\to x_\ell} \frac{d^{\eta_\ell}}{dz^{\eta_\ell}} z^{j+k-2} S_\ell(z) v_j v_k$$

$$= \sum_{\ell=1}^{\tilde{n}} \lim_{z\to x_\ell} \frac{d^{\eta_\ell}}{dz^{\eta_\ell}} S_\ell(z) \sum_{j=1}^{n_x}\sum_{k=1}^{n_x} z^{j+k-2} v_j v_k = \sum_{\ell=1}^{\tilde{n}} \lim_{z\to x_\ell} \frac{d^{\eta_\ell}}{dz^{\eta_\ell}} S_\ell(z) P_v(z)^2. \tag{6.9}$$

If

$$\tilde{n} + \sum_{j=1}^{\tilde{n}} \lfloor \eta_j/2 \rfloor \leq n_x - 1, \tag{6.10}$$

then we can take

$$P_v(z) = (z - x_1)^{1+\tilde{\eta}_1}(z - x_2)^{1+\tilde{\eta}_2} \cdots (z - x_{\tilde{n}})^{1+\tilde{\eta}_{\tilde{n}}}, \qquad \tilde{\eta}_j = \lfloor \eta_j/2 \rfloor.$$

Since $2(1 + \tilde{\eta}_\ell) = 2 + 2\lfloor \eta_\ell/2 \rfloor \geq 2 + 2(\eta_\ell/2 - 1) > \eta_\ell$ and

$$\left(\frac{d^\ell}{dz^\ell} f(z)(z - z^*)^k \right)\Bigg|_{z=z^*} = 0, \qquad 0 \leq \ell < k,$$

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

for all smooth enough $f(z)$, we get $\boldsymbol{v}^T A_1 R \boldsymbol{v} = 0$, which contradicts the positivity of $A_1 R$. Hence,

$$\tilde{n} + \sum_{j=1}^{\tilde{n}} \lfloor \eta_j/2 \rfloor > n_x - 1 = \tilde{n} + \sum_{\ell=1}^{\tilde{n}} \eta_\ell - 1.$$

Since for any integer $n > 0$ we have $\lfloor n/2 \rfloor \leq n - 1$, it follows that all $\eta_\ell = 0$ and $\tilde{n} = n_x$. Hence, if $A_1 R$ is positive definite, then $\{x_j\}$ are distinct.

To show the theorem it is now enough to show that, when $\{x_j\}$ are distinct, $A_1 R$ is positive if and only if the weights are positive. From (6.9) we then have

$$\boldsymbol{v}^T A_1 R \boldsymbol{v} = \sum_{\ell=1}^{n_x} S_\ell(x_\ell) P_v(x_\ell)^2 = \sum_{\ell=1}^{n_x} w_\ell P_v(x_\ell)^2.$$

Clearly, when all $w_\ell > 0$, this expression is positive for $\boldsymbol{v} \neq 0$, and $A_1 R$ is positive definite. To show the converse, we take $P_v(z)$ to be the Lagrange basis polynomials $L_j(z)$ of degree $n_x - 1$ defined as

$$L_j(x_i) = \begin{cases} 1, & i = j, \\ 0, & i \neq j. \end{cases}$$

If $A_1 R$ is positive, then

$$0 < \boldsymbol{v}^T A_1 R \boldsymbol{v} = \sum_{\ell=1}^{n_x} w_\ell L_j(x_\ell)^2 = w_j.$$

This can be done for each j, which concludes the proof. □

We can now relate our conclusions with those in Markov's theorem, Theorem 6.1. We consider all solutions to (1.5) instead of those given by the integral relation (6.1) with a piecewise continuous function $f(x)$. The extra condition (6.2) is then automatically satisfied, and we note that the positivity of $A_1 R$ guarantees a unique solution also in our space of density functions (1.6). We view this as a corollary of Theorems 4.1 and 6.4.

COROLLARY 6.5. *If there exists a solution to* (1.5)*, then the matrix in* (6.2) *is singular. When $n_x = n_y$, there is a unique solution to* (1.5) *of the form* (1.3) *if and only if $A_1 R$ is symmetric positive definite.*

Proof. We start by proving the singularity of (6.2). By (ii) in Proposition 4.6 a coefficient solution $\boldsymbol{c} = (c_0, \ldots, c_{n_x})^T = (c_0, \bar{\boldsymbol{c}}^T)^T$ satisfies $A\boldsymbol{c} = 0$. Since $A = (\boldsymbol{a}_0\ A_1)$, it remains to prove that $c_0 a_{K+1} + \boldsymbol{a}_0^T \bar{\boldsymbol{c}} = 0$. This was already proved in (5.2).

Next, we prove the "if" part of the second statement. If $A_1 R$ is symmetric positive definite, it is nonsingular, and by (i), (iii), and (v) in Theorem 4.1, the minimal degree solution exists and is unique and $x_j \neq y_i$ for all i, j. (If $x_j = 0$ for some j, then there is no zero y_i-value since $\mathrm{Deg}(q^*) = n$ by point (iii).) By Theorem 6.4 the corresponding branch values $\{x_j\}$ are distinct. It remains to show that, upon some reordering, the $\{x_j\}$ and $\{y_j\}$ are interlaced as in (1.3).

Order the x_j-values in an increasing sequence and let m_k be the number of y_j-values such that $y_j < x_k$. Clearly, m_k is increasing and $0 \leq m_k \leq n_y$. Moreover, $\mathrm{sgn}(q_r(x_k)) = (-1)^{n_y - m_k}$, and since $\lim_{z\to\infty} p_r'(z) > 0$, we also have $\mathrm{sgn}(p_r'(x_k)) = (-1)^{n_x - k}$. Hence, by also using the fact that $n_y = n_x$,

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

$$\mathrm{sgn}(w_k) = (-1)^{n_y - m_k + n_x - k} = (-1)^{m_k + k}.$$

We conclude that $m_k + k$ is even, which implies that m_k is, in fact, *strictly* increasing. Then, for $k = 1, \dots, n_x - 1$, we have $m_{k+1} \geq m_k + 1$ and

$$n_x \geq m_{n_x} \geq m_k + n_x - k \quad \Rightarrow \quad m_k \leq k.$$

Similarly, $m_k \geq m_1 + k - 1 \geq k - 1$, so $k - 1 \leq m_k \leq k$, and therefore

$$2k - 1 \leq m_k + k \leq 2k.$$

Finally, since $m_k + k$ is even, we must have $m_k = k$, which implies that the values are interlaced.

We now consider the "only if" part. If there is a solution of the form (1.3), then the $\{x_j\}$-values are obviously distinct and $m_k = k$. By Proposition 6.3 the weights are then given by (6.6) and they are positive since, as above, $\mathrm{sgn}(w_k) = (-1)^{m_k + k} = 1$. It follows from Theorem 6.4 that $A_1 R$ is positive definite. □

7. Outlook. Several interesting issues may be worth mentioning:

1. Computational complexity in a finite difference implementation: One can consult the article [14], where practical implementation issues and several examples of increasing complexity have been addressed in the context of geometric optics problems. In particular, comparisons with Lagrangian (ray-tracing) solutions are shown.
2. Extension to higher dimensions for the present problem: Nothing seems to exist in this direction at the time being; see, however, the last sections of [20] and the routines based on complex variables in [11, 9] for "shape from moments."
3. A very special case of the trigonometric moment problem can be solved by means of a slight variation of the algorithms presented here, in [14], and in section IV.A of [9]. That is to say, one tries to invert the following set of equations:

$$\sum_{j=0}^{n} \mu_j \exp(ik\lambda_j) = m_k, \qquad k = 0, \dots, n. \tag{7.1}$$

Let us state that in case the $n + 1$ real frequencies λ_j are known, the set of complex amplitudes μ_j are found by solving a Vandermonde system:

$$\begin{pmatrix} 1 & \cdots & 1 \\ \exp(i\lambda_0) & \cdots & \exp(i\lambda_n) \\ \vdots & & \vdots \\ \exp(in\lambda_0) & \cdots & \exp(in\lambda_n) \end{pmatrix} \begin{pmatrix} \mu_0 \\ \mu_1 \\ \vdots \\ \mu_n \end{pmatrix} = \begin{pmatrix} m_0 \\ m_1 \\ \vdots \\ m_n \end{pmatrix}.$$

The frequencies can be found through a byproduct of [9, 14] as we state now. Let us suppose n is odd (i.e., the number of equations is even); we form the two matrices

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

$$A_1 = \begin{pmatrix} m_0 & \cdots & m_{\frac{n-1}{2}} \\ \vdots & & \vdots \\ m_{\frac{n-1}{2}} & \cdots & m_{n-1} \end{pmatrix}, \qquad A_2 = \begin{pmatrix} m_1 & \cdots & m_{\frac{n+1}{2}} \\ \vdots & & \vdots \\ m_{\frac{n+1}{2}} & \cdots & m_n \end{pmatrix},$$

and then the frequencies can be obtained through a generalized eigenvalue problem, $A_1 \boldsymbol{v}_j = \lambda_j A_2 \boldsymbol{v}_j$, $j = 0, \ldots, n$. This kind of algorithm can be used to check the accuracy of the classical FFT and will be studied in a forthcoming article.

REFERENCES

[1] N. I. Akhiezer, *The Classical Moment Problem and Some Related Questions in Analysis*, Oliver and Boyd, Edinburgh, 1965.

[2] N. I. Akhiezer and M. G. Krein, *Some Questions in the Theory of Moments*, Transl. Math. Monogr. 2, AMS, Providence, RI, 1962.

[3] Y. Brenier, *Équations de moment et conditions d'entropie pour des modèles cinétiques* (French) [Moment equations and entropy conditions for kinetic models], in Séminaire sur les Équations aux Dérivées Partielles, 1994–1995, École Polytech., Palaiseau, France, 1995, Exp. No. XXII.

[4] Y. Brenier and L. Corrias, *A kinetic formulation for multibranch entropy solutions of scalar conservation laws*, Ann. Inst. H. Poincaré Anal. Non Linéaire, 15 (1998), pp. 169–190.

[5] R. E. Curto and L. A. Fialkow, *The truncated complex K-moment problem*, Trans. Amer. Math. Soc., 352 (2000), pp. 2825–2855.

[6] D. Czarkowski, D. V. Chudnovsky, G. V. Chudnovsky, and I. W. Selesnick, *Solving the optimal PWM problem for single-phase inverters*, IEEE Trans. Circuits Systems I Fund. Theory Appl., 49 (2002), pp. 465–475.

[7] P. Diaconis, *Application of the method of moments in probability and statistics*, in Moments in Mathematics (San Antonio, TX, 1987), Proc. Sympos. Appl. Math. 37, AMS, Providence RI, pp. 125–142.

[8] P. Diaconis and D. Friedman, *The Markov moment problem and de Finetti's theorem*, Math. Z., 247 (2004), pp. 183–199, 201–212.

[9] M. Elad, P. Milanfar, and G. H. Golub, *Shape from moments—an estimation theory perspective*, IEEE Trans. Signal Process., 52 (2004), pp. 1814–1829.

[10] F. Gamboa and L. V. Lozada-Chang, *Large-deviations for random power moment problems*, Ann. Probab., 32 (2004), pp. 2819–2837.

[11] G. H. Golub, P. Milanfar, and J. Varah, *A stable numerical method for inverting shape from moments*, SIAM J. Sci. Comput., 21 (1999), pp. 1222–1243.

[12] L. Gosse, *Using K-branch entropy solutions for multivalued geometric optics computations*, J. Comput. Phys., 180 (2002), pp. 155–182.

[13] L. Gosse, S. Jin, and X. Li, *Two moment systems for computing multiphase semiclassical limits of the Schrödinger equation*, Math. Models Methods Appl. Sci., 13 (2003), pp. 1689–1723.

[14] L. Gosse and O. Runborg, *Finite moment problems and applications to multiphase computations in geometric optics*, Commun. Math. Sci., 3 (2005), pp. 373–392.

[15] L. Gosse and O. Runborg, *Resolution of the finite Markov moment problem*, C. R. Math. Acad. Sci. Paris, 341 (2005), pp. 775–780.

[16] V. I. Korobov and G. M. Sklyar, *Time-optimality and the power moment problem*, Mat. Sb. (N.S.), 134 (176) (1987), pp. 186–206, 287 (in Russian); translation in Math. USSR-Sb., 62 (1989), pp. 185–206.

[17] M. G. Krein and A. A. Nudel'man, *The Markov Moment Problem and Extremal Problems*, Transl. Math. Monogr. 50, AMS, Providence, RI, 1977.

[18] A. S. Lewis, *Superresolution in the Markov moment problem*, J. Math. Anal. Appl., 197 (1996), pp. 774–780.

[19] D. T. Norris, *Optimal Solutions to the L^∞ Moment Problem with Lattice Bounds*, Ph.D. thesis, University of Colorado, Boulder, CO, 2002; available online from http://math.colorado.edu/~norrisdt/dougthesis.ps.

[20] M. Putinar, *A renormalized Riesz potential and applications*, in Advances in Constructive Approximation, M. Neamty and E. Saff, eds., Nashboro Press, Brentwood, TN, 2004, pp. 433–465.

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

[21] O. Runborg, *Some new results in multiphase geometrical optics*, M2AN Math. Model. Numer. Anal., 34 (2000), pp. 1203–1231.

[22] M. I. Sezan and H. Stark, *Incorporation of a-priori moment information into signal recovery and synthesis problems*, J. Math. Anal. Appl., 122 (1987), pp. 172–186.

[23] B. Simon, *The classical moment problem as a self-adjoint finite difference operator*, Adv. Math., 137 (1998), pp. 82–203.

[24] G. M. Sklyar and L. V. Fardigola, *The Markov power moment problem in problems of controllability and frequency extinguishing for the wave equation on a half-axis*, J. Math. Anal. Appl., 276 (2002), pp. 109–134.

[25] G. Talenti, *Recovering a function from a finite number of moments*, Inverse Problems, 3 (1987), pp. 501–517.

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.